

doi:10.11918/j.issn.0367-6234.2016.05.020

一种快速有效的遥感图像场景分类特征

徐 侃¹, 胡 凡², 李 陶¹

(1. 武汉大学 卫星导航定位技术研究中心, 430079 武汉; 2 武汉大学 电子信息学院, 430079 武汉)

摘 要: 针对在遥感大数据中如何进行快速有效的场景分类,提出一种图像特征的构建方法. 基于非监督学习进行快速二值编码,首先对图像局部训练样本利用非监督学习获取相应滤波器组,然后再使用二值化哈希编码方法对场景单元特征图进行量化,最后统计得到场景全局特征. 实验结果表明,该特征描述结合了滤波器组和二进制特征描述子的优点,在保证较高分类精度的前提下,能够大幅度提升计算效率,具有较好的鲁棒性.

关键词: 遥感图像; 场景分类; 二值编码; 特征表达

中图分类号: P237.4

文献标志码: A

文章编号: 0367-6234(2016)05-0122-05

A feature for fast and effective scene classification of remote sensing image

XU Kan¹, HU Fan², LI Tao¹

(1. GNSS Research Center, Wuhan University, 430079 Wuhan, China;
2. School of Electronic Information, Wuhan University, 430079 Wuhan, China)

Abstract: For this purpose of the rapid and effective scene interpretation for remote sensing big data, a novel method for feature representation is presented. First of all, the corresponding filter banks of local training samples in image are obtained by unsupervised learning. The feature map of scene units is then quantized based on binary hashing coding. Finally, the global feature of scene is obtained from statistical results. The experimental results demonstrate that, combining with both the advantages of filter bank and binary feature descriptor, the proposed feature can greatly enhance the computational efficiency on the assumption of ensuring high accuracy, which has shown its robustness.

Keywords: remote sensing image; scene classification; binary coding; feature representation

随着遥感技术的飞速发展,遥感影像已成为各部门及机构直接有效获取对地观测信息的重要技术手段^[1]. 面对海量的高分辨率遥感影像数据,快速识别和获取不同专题信息通常直接取决于对遥感图像场景的内容解译. 而遥感影像场景分类则是对整幅图像最直观的理解,分类的准确性直接决定了图像场景解译质量的高低,分类结果的准确性与所关注地物的有效提取,将成为相关部门的重要参考依据,为进一步决策或规划提供有力技术支撑. 面对源源不断所获取的海量遥感数据与信息,如何针对其进行高效处理,迅速提取用户密切关注的重点区域,一直以来都是备受关注的研究热点. 因此,作为

遥感图像场景与地物分类的先决条件和重要步骤,研究其特征提取和表达方法具有重要意义.

目前应用于遥感图像分类的特征多种多样,其本质为对数字信号进行处理分析^[2]. 而使用滤波器组的方法在信号处理分析领域中是长期备受关注的焦点. 以纹理图像识别为例,通常先将纹理图像与已设定好的一组滤波器进行卷积运算,连接各滤波器的响应;然后利用量化方法构造纹理基元;最后统计图像中不同纹理基元频率直方图,将其作为纹理图像特征^[3-5]. 尽管这种方法在纹理识别中非常有效,识别精度也比较高,但是由于构造纹理基元时需要大量训练数据,导致量化方法(实际情况一般采用 K-means 算法)的训练过程十分缓慢,计算代价太大. 另一方面,由于二进制局部特征(binary local feature)描述算子具有计算简便快速且便于存储的特点,使得其在图像特征表达研究领域得到了广泛应用. 但基于该框架所构建的特征鲁棒性和判别性较弱,往往会影响图像分类精度^[6-10].

收稿日期: 2015-04-13.

基金项目: 国家自然科学基金项目(41274048),武汉大学自主科研项目(2042014kf0082).

作者简介: 徐 侃(1983—),男,博士.

通信作者: 徐 侃, jcarloswhu@msn.com.

针对上述方法所存在的不足,本文结合滤波器组和二进制特征描述算子的优点,尤其针对于大尺寸高分辨率遥感影像,提出一种基于快速二值编码特征表达方法,应用于高分辨率遥感影像场景分类.本文方法采用非监督方法学习场景单元中的局部图像块,通过滤波器组与场景单元的卷积响应进行二值编码,获取各场景图像的全局描述,不仅具有高分类精度,同时还能大幅提高计算效率.

1 基于快速二值编码的特征表达

利用非监督学习算法训练场景单元中的局部图像块获得滤波器组,并将滤波器组与场景单元的卷积响应进行二进制编码,获得大幅遥感场景单元的全局特征描述.

1.1 遥感影像场景划分

对大场景遥感影像进行场景分类,首先需要定义场景单元与场景类别数.本文在大幅遥感影像中选择合适尺寸的矩形区域作为场景单元,采用均匀网格进行划分,每个子网格代表一个场景单元且相邻场景单元间无重叠.

1.2 局部图像块训练样本提取

分类框架中需要对获得的场景单元进行学习,得到数据驱动型滤波器组,滤波器组需要从场景单元中大量无标号的局部图像块(patch)中自动学习得到.针对上一步骤获得的初始化无标号场景单元,从各场景单元中随机抽取若干尺寸相同的patch构成训练样本,将各patch在像素级(pixel)空间上展开为一个向量,该向量中的元素为patch的像素值,并对其进行均值方差归一化操作:

$$\mathbf{x}^{(i)} = (\mathbf{p}^{(i)} - \text{mean}(\mathbf{p}^{(i)})) / \text{std}(\mathbf{p}^{(i)}). \quad (1)$$

式中: $\mathbf{p}^{(i)}$ 表示第*i*个patch向量, $\mathbf{x}^{(i)}$ 表示均值方差归一化后第*i*个patch向量, $\text{mean}(\mathbf{p}^{(i)})$ 和 $\text{std}(\mathbf{p}^{(i)})$ 分别表示第*i*个patch向量中元素的均值和方差.将归一化后的patch向量 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(M)}$ 构建成向量矩阵 $\mathbf{X} = [\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(M)}], \mathbf{x}^{(i)} \in \mathbf{R}^{n \times 1}, \mathbf{R}^n$ 表示*n*维实空间,且 $n = r \times r, r \times r$ 为patch尺寸.*M*表示patch训练样本个数;*i*为patch训练样本序号,其取值为1,2,...,*M*.

1.3 利用非监督学习获得滤波器组

在获得了均值方差归一化的patch向量矩阵 $\mathbf{X}$ 后,需要选择合适的非监督学习方法对其进行训练,由此获得滤波器组.具体实施方式中,可以采用多种不同的非监督学习方法来获得滤波器组,在本文的实验中以稀疏编码^[11]为例,具体说明如何从patch向量矩阵 $\mathbf{X}$ 中学习得到滤波器组.

输入patch向量矩阵 $\mathbf{X}$,将采用稀疏编码法学习

得到的字典作为滤波器组.传统的稀疏编码方法在进行非监督训练时,视觉字典和与之对应的稀疏向量通过求解如下优化问题来获取:

$$\begin{aligned} \min_{\mathbf{W}, \mathbf{z}^{(i)}} \quad & \|\mathbf{W}\mathbf{z}^{(i)} - \mathbf{x}^{(i)}\|_2^2 + \lambda \|\mathbf{z}^{(i)}\|_1 \\ \text{s.t.} \quad & \|\mathbf{W}^{(k)}\|_2^2 \leq 1, \forall k \end{aligned}$$

式中: $\mathbf{W}$ 表示由稀疏编码法学习得到的字典; $\mathbf{W}^{(k)}$ 表示字典 $\mathbf{W}$ 中第*k*个码本, $\mathbf{W}^{(k)} \in \mathbf{R}^{n \times 1}; \mathbf{z}^{(i)}$ 表示第*i*个patch训练样本的稀疏向量; λ 为惩罚权重因子,用来控制稀疏向量 $\mathbf{z}^{(i)}$ 的稀疏度,即稀疏向量中零元素个数.在完成稀疏编码后,将字典 $\mathbf{W}$ 作为滤波器组, $\mathbf{W}^{(k)}$ 表示第*k*个滤波器.

除了上述方法,还可以采用其他非监督学习方法,例如:K均值(K-Means)、主成分分析PCA(principal component analysis)、局部保持映射LPP(locality preserving projections)、独立成分分析ICA(independent component analysis)、非负矩阵分解NMF(non-negative matrix factorization)、随机映射RP(random projection)等,同样也可以获得滤波器组.

1.4 场景单元的快速二值编码

通过非监督学习,可以从大量无标号的patch训练样本中学习得到滤波器组 $\mathbf{W}$,利用滤波器组中 $\mathbf{W}$ 各滤波器分别与场景单元进行卷积计算,获得各场景单元的滤波器响应图.在做滤波器卷积前,需重新调整滤波器组 $\mathbf{W}$ 中各滤波器 $\mathbf{W}^{(k)}$ 尺寸,使得滤波器 $\mathbf{W}^{(k)}$ 与patch尺寸相同.本文采用函数 $\text{resize}(\cdot): \mathbf{R}^{n \times 1} \mapsto \mathbf{R}^{r \times r}$ 调整滤波器尺寸,尺寸调整后滤波器为 $w^{(k)} = \text{resize}(\mathbf{W}^{(k)}), w^{(k)} \in \mathbf{R}^{r \times r}$.完成滤波器尺寸调整后,滤波器组 $\mathbf{W}$ 就变为 $\mathbf{W}^w \in \mathbf{R}^{r \times r \times L}, L$ 表示滤波器个数.利用调整尺寸的滤波器组 $\mathbf{W}^w$ 与各场景单元分别做卷积,获得*N*个滤波器响应图,该过程可采用如下公式表示:

$$f_k^{(i)} = S^{(i)} * w^{(k)}.$$

式中: $S^{(i)}$ 表示第*i*幅场景单元, $i = 1, 2, \dots, N, N$ 为遥感影像中场景单元总数; $f_k^{(i)}$ 表示第*k*个滤波器 $w^{(k)}$ 与场景单元 $S^{(i)}$ 卷积获得的第*k*幅滤波器响应图,即特征图; $k = 1, 2, \dots, L, L$ 表示滤波器组中滤波器个数;“*”表示二维卷积操作(滤波器沿各像素滑动均进行相同操作).

为得到各场景单元的全局特征描述,需要将各场景单元的*L*个特征图融合,本文提出了一种二值编码法来实现场景单元特征图的融合.为了使二值编码过程更简洁,免除由于特征图在量级上的差异所采取的尺度归一化等不必要的操作,需要选取合适阈值(本文中设定为0).将特征图上各位置滤波器响应值与阈值作比较,大于阈值的位置编码为1,

小于阈值的位置编码为 0, 编码过程可用如下公式表示:

$$F_k^{(i)} = H(f_k^{(i)}).$$
$$H(x) = \begin{cases} 1, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

式中: $F_k^{(i)}$ 为特征图 $f_k^{(i)}$ 对应的二值编码图, $H(x)$ 为单位阶跃响应函数. 获得了场景单元的 L 个二值编码图后, 将场景单元的所有二值编码图中对应位置上的编码组合成 L 位的二进制数, 并将该二进制数转化成十进制数, 从而获得融合的十进制编码图, 整个过程用下式表示:

$$T^{(i)}(a,b) = \sum_{k=1}^L 2^{k-1} \cdot F_k^{(i)}(a,b).$$

式中: $T^{(i)}$ 表示第 i 幅场景单元的十进制编码图, (a,b) 为编码图中各位置坐标. 最后获得 $T^{(i)}$ 在非负整数区间 $[0, 2^L - 1]$ 上的直方图, 记为 $Y^{(i)}$, 并将 $Y^{(i)}$ 作为第 i 幅场景单元的全局特征描述. 滤波器组中滤波器的个数决定了十进制响应值的范围, 在场景单元十进制编码图中, 所有十进制响应值的频率直方图则作为该场景单元的全局特征描述.

2 实验结果

2.1 实验数据集

资源三号测绘卫星(ZY-3)集测绘和资源调查功能于一体, 是我国第一颗自主的民用高分辨率立体测绘卫星, 所配备的三线阵测绘相机和多光谱相机等有效载荷可以长期连续且稳定地获取全色影像、多光谱影像以及辅助数据, 能够为国土资源、农业、林业等多项领域提供服务. 资源三号测绘卫星影像所对应参数如表 1 所示. 为了验证本文方法的有效性, 选取资源三号数据中的 2.1 m 分辨率全色影像, 其覆盖范围涵括湖北省某输电线路走廊区域, 主要包括的地物类型有: 房屋、农田、水域、植被.

表 1 资源三号卫星影像参数

数据类型	分辨率/m	幅宽/km	波长/nm
星下点全色	2.1	50	—
前、后视全色	3.6	50	—
星下点多光谱	5.8	52	蓝:450–520
			绿:520–590
			红:630–690
			近红外:770–890

2.2 分类器

采用 SVM 分类器, 以最大分类间隔为准则, 通过寻找最优分类超平面来对样本进行空间划分. SVM 分类器的训练过程最终转化为解决如下优化

问题:

$$\min \frac{1}{2} \| \boldsymbol{w} \|^2 + C \sum_{i=1}^n \xi_i.$$
$$\text{s.t.} \quad y_i(\boldsymbol{w}^T \boldsymbol{\varphi}(x_i) + b) \geq 1 - \xi_i, i = 1, \cdots, n;$$
$$\xi_i \geq 0, i = 1, \cdots, n.$$

式中: $\boldsymbol{w}$ 和 b 为定义分类超平面 $\boldsymbol{w}^T \boldsymbol{\varphi}(x) + b = 0$ 的参数, 常量 C 用于控制目标函数中两项间的权重, ξ_i 表示松弛变量, $\boldsymbol{\varphi}(x)$ 表示训练样本 x 的非线性映射, y_i 表示第 i 个训练样本的类别标号, n 为训练样本个数.

2.3 实验流程

实验流程如图 1 所示. 训练数据集从输电线路走廊区域附近范围中选取, 每一类地物分别选取 10、20、30、40、50 幅训练图像作为训练样本, 大小为 100×100 像素. 所采用的测试图像(输电线路走廊区域缓冲区)大小为 $2\,000 \times 4\,000$ 像素, 在分类过程中被均匀划分为 800 个 100×100 像素的 patch. 在经过所构建的快速二值编码特征表达后, 输入到 SVM 分类器完成分类. 实验均建立在随机抽取训练与测试图像样本上进行, 计算每次实验中所有类别平均分类准确率, 每组实验重复 5 次, 将所得的平均准确率与标准偏差作为最终结果.

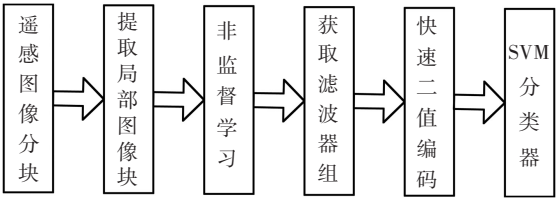


图 1 分类实验流程示意

2.4 分类结果

分类结果如图 2 所示. 在人工标注结果中, 不属于任何一类的被标注为空. 本文采用基于“词袋”BOW(bag of words)表示方法, 采用尺度不变特征变换(SIFT)^[2]进行分类性能比较. 其中, 对于 SIFT 特征使用 K 均值聚类进行聚类运算(聚类中心个数设为 300), 进而将其量化为视觉词汇表. 与之对应的分类结果如图 2 中第 2 列所示. 从图中可以看出, 与本文方法中所采用的特征相比, 基于 BOW 的特征表述容易将一部分水域和植被错分为农田, 这是由于水域和农田在结构类型上具有一定程度的相似性, 而在农田的周围所分布的植物在 BOW 表述方法下对分类起到了干扰作用. 由此可见, 本文所提出的特征描述方法具有更好的判别性. 基于本文特征表达所得分类结果的混淆矩阵如图 3 所示, 可见对于水域的错分现象明显减少, 而错分现象主要出现在农田与植被之间. 这是由于在灰度图像中农田

与植被的纹理类型比较接近,而本文所构建的二值化特征主要侧重于对地物纹理特性的描述,从而容

易导致混淆.

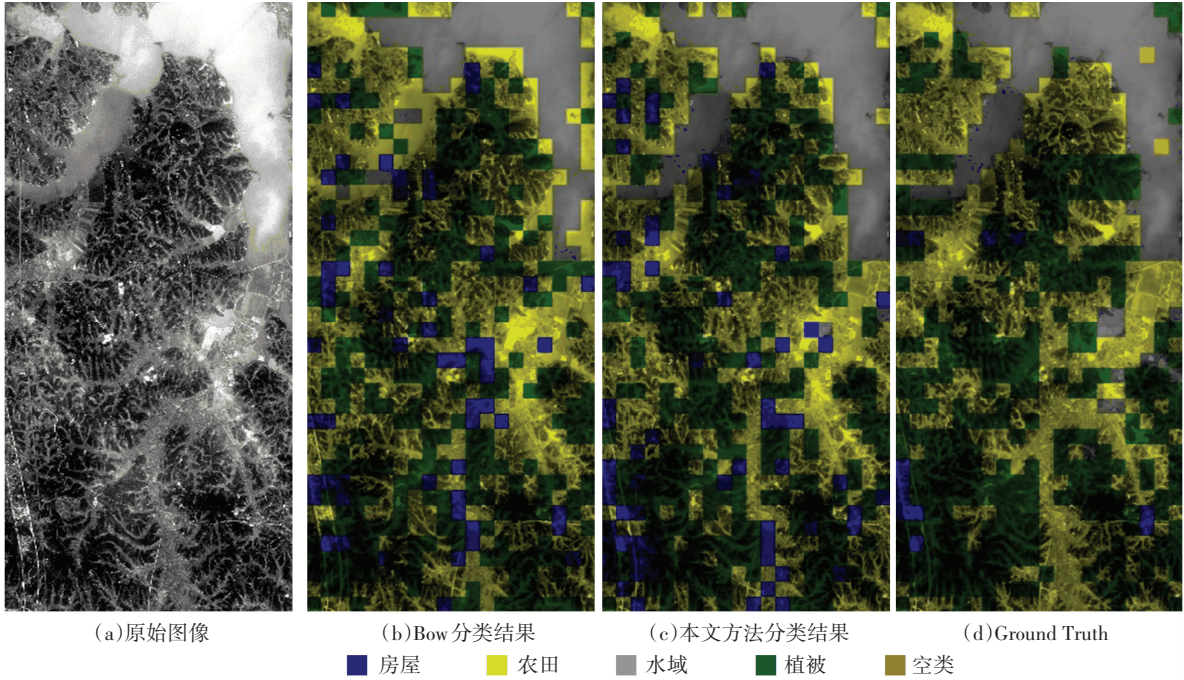


图 2 分类结果示意

近年来,输电线路因野外山火而引发的跳闸停电故障不断增多,给电网的安全可靠运行带来了巨大隐患. 因此,输电线路中的火险高发地带属于电网运维管理工作中所关注的重点区域. 由本文的分类结果可以从线路走廊区域中迅速剔除不易发生火险的区域(水域与房屋),而对于容易发生火险的区域(植被与农田)则属于需要在输电线路运维中着重关注的对象.

能够将计算时间从数小时大幅度缩减至数十秒,计算速率得以有效提高.

表 2 不同训练样本下分类结果比较

特征种类	训练图像数量				
	10	20	30	40	50
SIFT	40.2±8.8	43.8±8.2	46.8±7.8	50.3±8.2	54.4±7.4
FBC	49.8±8.4	53.6±6.8	57.3±6.4	60.2±7.6	64.0±6.8

表 3 分类精度及计算速率比较(训练样本=50 幅)

特征表述方法	分类精度/%	计算时间/s
BOW(SIFT)	54.4	>7 200
FBC	64.0	12

3 结 论

1) 针对从遥感图像所覆盖的输电线路走廊区域中如何快速有效提取地物类型这一问题,通过对图像局部训练样本获取滤波器组,采用二值化哈希编码对特征图进行量化,构建一种快速二值编码的特征,应用于高分辨率遥感图像中输电线路走廊区域的地物分类.

2) 基于“资源三号”卫星影像的实验结果表明,本文所构建的特征在训练样本有限的情况下具有较高判别性,能够在保证较高分类精度的前提下,大幅提高计算效率. 分类结果具有较好鲁棒性,能够对

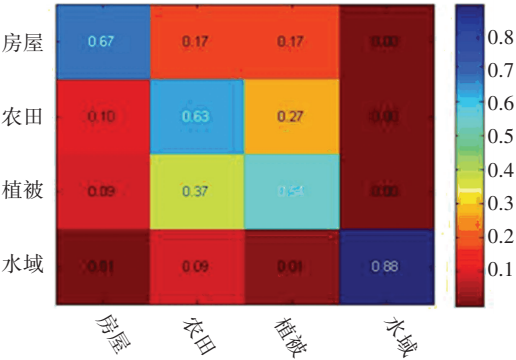


图 3 基于本文方法所得分类混淆矩阵

与此同时,使用 SIFT 特征,基于 BOW 的特征表示形式与本文方法所构建的特征在分类精度和计算效率方面分别进行了比较,对比结果如表 2 和表 3 所示. 可以看出,在训练样本相同的情况下,本文所构建的特征所对应分类结果具有更高分类精度,在训练样本较小的情况下仍具有较好的鲁棒性,而且

输电线路走廊区域中较高风险区域进行快速有效提取,可以为输电线路的运维工作提供参考依据.

参考文献

[1] 宋维静,刘鹏,王力哲,等. 遥感大数据的智能处理:现状与挑战[J]. 工程研究——跨学科视野中的工程, 2014, 6(3):259-265.

[2] 徐侃,陈丽君,杨文,等. 利用特征选择的遥感图像场景分类[J]. 哈尔滨工业大学学报, 2011, 43(9): 117-121.

[3] LEUNG T, MALIK J. Representing and recognizing the visual appearance of materials using three-dimensional textons [J]. International Journal of Computer Vision, 2001, 43(1): 29-44.

[4] VARMA M, ZISSERMAN A. A statistical approach to texture classification from single images [J]. International Journal of Computer Vision, 2005, 62(1):61-81.

[5] LIU L, FIEGUTH P W. Texture classification from random features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(3): 574-586.

[6] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern

Analysis and Machine Intelligence, 2002, 24(7): 971-987.

[7] AHONEN T, RAHTU E, OJANSIVU V, et al. Recognition of blurred faces using local phase quantization [C]//19th International Conference in Pattern Recognition (ICPR). Florida:IEEE Computer Sciety, 2008:1-4.

[8] CALONDER M, LEPETIT V, STRECHA C, et al. Brief: binary robust independent elementary features [C]//European Conference on Computer Vision(ECCV). Grete: Springer, 2010: 778-792.

[9] LEUTENEGGER S, CHLI M, SIEGWART R Y. BRISK: binary robust invariant scalable keypoints [C]//IEEE International Conference on Computer Vision (ICCV). Barcelona: IEEE, 2011: 2548-2555.

[10] ALAHI A, ORTIZ R, VANDERGHEYNST P. Freak: fast retina keypoint[C]//IEEE Conference in Computer Vision and Pattern Recognition (CVPR).Providence, RI: IEEE, 2012: 510-517.

[11] MAIRAL J, BACH F, PONCE J,et al. Online dictionary learning for sparse coding[C]//International Conference on Machine Learning (ICML). Montreal: International Machine on Learning Society, 2009: 689-696.

(编辑 王小唯 苗秀芝)